基于特征重标定生成对抗网络的图像分类算法 *

姜代红 [1]，张三友 [2, 3]，刘其开 [3]

(1. 徐州工程学院 信电工程学院，江苏 徐州 221008; 2. 苏州市吴江区公安局 科技信息化大队，江苏 苏州 215200; 3. 中国矿业大学 信息与控制工程学院，江苏 徐州 221008)

摘　要：针对传统鉴别器的损失策略和结构难以提取到更抽象以及任务相关的鲁棒性特征，从而导致半监督图像分类表现不足，提出了基于特征重标定的生成对抗网络。为了学习到任务相关的特征，在现有半监督 GAN 的基础上，为鉴别器引入模型在不同状态下的无监督均方差损失正则项，对训练样本中两个分支的同一输入对应得到的不同输出进行参数惩罚，从而指导特征重标定的优化方向。此外，在鉴别器中加入压缩激活模块来优化传统鉴别器的卷积池化结构。该模块自动学习每一个特征通道的重要程度，能够提取任务相关的特征抑制任务无关的特征，实现特征的重标定功能，从而提高半监督图像分类的表现。

关键词：生成对抗网络；图像分类；特征重标定；深度学习

中图分类号：TP391.41　　doi: 10.19734/j.issn.1001-3695.2018.08.0668

Image classification algorithm based on feature recalibration GAN

Jiang Daihong[1], Zhang Sanyou[2, 3], Liu Qikai[3]

(1. *School of Information & Electronic Engineering*, *Xuzhou Institute of Technology*, *Xuzhou Jiangsu* 221008, *China*; 2. *Dept. of Science & Technology*, *Suzhou Wujiang District Public Security Bureau*, *Suzhou* Jiangsu 215200, *China*; 3. *School of Information & Electrical Engineering*, *China University of Mining & Technology*, *Xuzhou Jiangsu* 221008, *China*)

Abstract: In view of the loss strategy and structure of traditional discriminator is difficult to extract more **Abstract:** and task related robustness features, which leads to the shortage of semi supervised image classification, and proposes a generation countermeasure network based on feature re calibration. In order to learn the related features of the task, on the basis of the existing semi supervised GAN, the unsupervised canonical loss regular term of the model under different states is introduced to the discriminator, and the different output of the same input corresponding to the two branches of the training sample is punished to guide the optimization direction of the feature re calibration. In addition, the compression activation module is added to the discriminator to optimize the convolution pool structure of the traditional discriminator. The module automatically learns the importance of each characteristic channel, and can extract the features related to the task to suppress the unrelated features of the task, and realize the recalibration function of the feature, thus improving the performance of the semi supervised image classification.

Key words: generative adversarial network; image classification; feature recalibration; deep learning

0　引言

图像分类是计算机视觉中最为主要的任务之一，而半监督图像分类算法能够在有限带标签样本的前提下较好地解决分类性能不足的问题。这几年卷积神经网络（CNN）对图像的特征提取使得视觉任务达到了一个新的高度。其中 Kingma 等人[1]提出了灵活且可伸缩的深度生成网络算法，采用丰富的参数密度估计来建模以及变分推理方式来训练，包括隐含特征判别模型（M1）和生成半监督模型（M2），以及堆叠而成的半监督模型（M1+M2），该模型取得较好的分类表现。Maaløe 等人[2]对深度生成网络模型进行了扩展，提出了辅助深度生成网络（ADGM），通过引入辅助的隐变量，改进变分近似算法，使得变分近似更加具有表现力。Li 等人[3]提出了一种最大间隔的深度生成模型（MMDGM），在保留强大的生成能力的同时，提高生成模型的预测性能，将一个边缘分类器的预测结果作为缺失的标签。Rasmus 等[4]提出了基于自动编码器的阶梯网络（ladder network），编码网络用于监督学习，解码网络的每一层与编码器相对应，用于无监督学习训练，形成阶梯，将每一层的重构损失作为总的损失函数来指导参数优化的方向。Laine 等人[5]提出简单高效的深度神经模型（Π-model 和 Temporal ensembing model），借鉴自学习思想，将无标签数据样本的预测结果来指导模型参数的优化方向，即在不同数据扩展的条件下为监督损失引入了无监督正则项，模型在半监督图像分类中取得较好的效果。

GAN 神经网络模型，在半监督学习上有着良好的表现。Miyato 等人[6]提出了一种基于虚拟对抗训练（VAT）的正则化方法:局部分布平滑（LDS）。Salimans 等人[7]通过一些技巧改善了常规 GAN 的性能，将鉴别器的二分类的概率输出扩展到多类别概率输出，即总类别为真实样本所属类别加生成样本类别，从而实现预测。Springenberg 等人[8]提出的

收稿日期：2018-08-06；**修回日期：**2018-09-28　**基金项目：**国家自然科学基金资助项目（61379100）；江苏省高等学校自然科学研究重大项目（18KJA520012）；徐州市科技计划项目基金资助（KC16SQ178）

作者简介：姜代红（1969-），女，江苏徐州人，教授，博士，主要研究方向为图像处理、智能计算（daihongjiang@163.com）；张三友（1989-），男，江西萍乡人，博士研究生，主要研究方向为视频警务、智能计算；刘其开（1992-），男，江苏徐州人，硕士，主要研究方向为深度学习、计算机视觉.

Cat-GAN 在鉴别器中引进适当的目标函数来权衡输入样本与对应预测类别的互信息，通过最小化真实数据样本类别的条件交叉熵和最大化生成数据类别的条件交叉熵来训练分类器，提高分类器的表现。Li 等人[9]提出了 triple GAN，通过引入额外的分类器，改善了 GAN 在半监督学习上的生成器和鉴别器在训练时无法同时达到最优的问题且生成器能够学习到样本的语义特征。Dumoulin 等人[10]提出 adversarially learning（ALI）推理模型，该模型引入一个推理器。它将数据空间中的训练样本映射到隐变量空间，输出训练样本与隐变量的联合分布；而生成器输出隐变量与生成样本的联合分布。

模型的深度及空间维度是提升 CNN 的性能两个重要方面，Jaderberg 等人[11]将 Attention 机制引入到空间维度上，允许在网络中对数据进行空间操作，使神经网络能够主动地在空间上进行特征映射。除了在空间维度上能够提升网络的性能，特征通道也是一个重要的突破点，Hu 等人[12]提出了压缩激活模块(squeeze-and-excitation networks, SENet)。该模块可以显式地构建特征通道之间的相互依赖关系，提取出有用的特征，抑制那些任务无关的特征。

因此，为了提高模型的泛化能力以及半监督分类的表现，提出了基于特征重标定的生成对抗网络，为鉴别器引入模型在不同状态下的无监督均方差损失，对同一输入所预测的不同输出进行参数惩罚。与此同时，在鉴别器中加入了 SENet 模块，自动学习每一个特征通道的重要程度，进行特征重标定操作，从而提高半监督学习表现。

1　基于特征重标定的生成对抗网络结构

基于特征重标定的生成对抗网络结构如图 1 所示，生成器的输入是维度为 100 的随机噪声 z ，$z \in R^{100}$ ，首先映射到维度为 1024×16 的高维空间，然后重新调整到通道数为 1024、大小为 4×4 的特征图，经过若干层的反卷积（deconvolutional network，DN）运算，每一次反卷积操作特征图的大小变为原来的两倍，最终得到与输入图像尺寸相同的生成样本。DN 是由 Zeiler 等人[13]提出的一种中低层特征学习的无监督算法，通过学习稀疏编码的参数来构建输入图像的中低层特征图。本质上学习输入图像潜在空间的特征图和卷积核对应的映射矩阵，然后通过学习到的卷积核重构出输入图像。

鉴别器的输入有生成样本和真实样本，真实样本中主要为无标签样本和有限的带标签样本。带标签样本仅参与鉴别器的监督损失，无标签样本和生成样本参与无监督对抗损失和无监督均方差损失。首先输入样本通过若干个的卷积与池化操作之后，得到一层特征图 U，其中 $U \in \mathrm{R}^{C \times H \times W}$ ，随后紧接一个 SENet 模块对特征进行标定，得到重标定的特征图 U′，其中 $\tilde{x} = G(z)$ 。为了增加网络的非线性拟合能力，引入 NIN 网络，其本质是进行一个 1×1 的卷积，为了减少模型的参数，防止过拟合，使用全局平均池化来代替全连接层，最终加权得到总的损失。

2　基于特征重标定生成对抗网络的图像分类算法

本文主要是从模型的集成思想和模型的结构出发，搭建了基于特征重标定的生成对抗网络。在训练过程中，由于网络模型 dropout 正则项以及随机噪声的存在，尽管在同一个迭代训练阶段且输入样本相同，但也会导致网络输出发生改变。因此，在同一个批迭代次数中，对同一数据进行随机处理，其输出同样会发生变化。为了指导模型参数优化的方向，希望同一个输入得到的输出尽可能相同，即对预测分类而言，归属的某一类的概率尽可能相近。

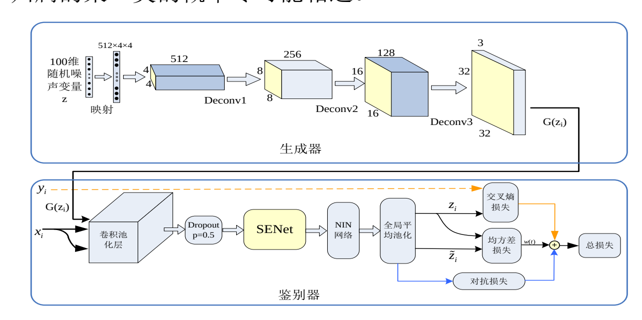


图 1　FR-GAN 的框架结构

Fig. 1　Frame structure of FR-GAN

2.1　算法原理

假设 $X=\{x^1,x^2,\cdots x^m\}$ 表示 m 个真实样本的集合， $z \in \mathbb{R}^{100}$ 表示维度为 100 的随机噪声向量， $\tilde{x} = G(z)$ 表示生成样本，则鉴别器的输出代表其输入样本属于真实样本的概率，表示为

$$p(y=1|x,D)=\frac{1}{1+e^{-D(x)}} \tag{1}$$

则常规 GAN 总的目标函数表示为

$$\min_G \max_D E_{x\sim X}[\log(p(y=1|x,D))]+E_{z\sim p_z}[\log(1-p(y=1|x,D))] \tag{2}$$

生成对抗网络的半监督学习主要是通过鉴别器来实现的，不仅要考虑到输入样本属于真实样本的概率，还要考虑带标签的输入样本对标签类别的概率。假设给定标签类别共有 K 类，样本 x 经过特征提取得到一个 K 维的特征向量 logits= $\{l_1,l_2,\cdots l_K\}$ ，则该样本对第 j 类的概率表示为

$$p_{\mathrm{model}}(y=j|x)=\frac{\exp(l_j)}{\sum_{k=1}^{K}\exp(l_k)} \tag{3}$$

在监督学习算法中，目标函数就是最小化样本标签与模型预测的概率值的交叉熵，为了获得 GAN 的监督损失，将输入到鉴别器的生成样本所属的类别定义为第 K+1 类，则来自生成样本的概率表示为 $p_{\mathrm{model}}(y=K+1|x)$ ，即对应于 $1-D(x)$ ，假设鉴别的训练样本一半来自真实样本，另一半来自生成样本，则半监督生成对抗网络鉴别器 D 的损失函数可表示为

$$\begin{aligned} L &= -E_{x,y\sim p_{data(x,y)}}[\log p_{\mathrm{model}}(y|x)]-E_{x\sim G}[\log p_{\mathrm{model}}(y=K+1|x)] \\ &= L_{\mathrm{supervised}}+L_{\mathrm{unsupervised}} \end{aligned} \tag{4}$$

带标签的训练样本贡献于监督损失为

$$L_{\mathrm{supervised}}=-E_{x,y\sim p_{data(x,y)}}[\log p_{\mathrm{model}}(y|x,y<K+1)] \tag{5}$$

无标签训练样本贡献于无监督对抗损失。令 $D(x)=1-p_{\mathrm{model}}(y=K+1|x)$ ，则

$$L_{\mathrm{unsupervised}}=-E_{x\sim p_{data}(x)}[\log(1-D(x))]-E_{z\sim p_{noise}}[\log(1-D(G(z)))] \tag{6}$$

其中： $D(x)=\frac{Z(x)}{Z(x)+1}$ ， $Z(x)=\sum_{k=1}^{K}\exp[l_k(x)]$ 。

PR-GAN 针对半监督分类任务，为鉴别器引入模型在不同状态下的无监督特征均方差损失，对真实样本两个分支的同一输入而得到的不同输出进行惩罚，该正则项能够学习任务相关的特征，提高半监督分类的精度。假设 $\tilde{X}=\{\tilde{x}^1,\tilde{x}^2,\cdots\tilde{x}^m\}$ 表示的是第二分支无标签的训练样本， $\tilde{D}(*)$ 表示的是鉴别器同一训练阶段下的不同状态，通过随机预处理函数 $f(*)$ 得到输入对应的输出为

$$\tilde{z}^{(i)}=\tilde{D}_\theta(f(\tilde{x}^{(i)})) \tag{7}$$

同理第一分支下的输入对应的输出为

$$z^{(i)}=D_\theta(f(x^{(i)})) \tag{8}$$

则无监督特征均方差正则项表示如下：

$$loss_{unsup}=w(t)\frac{1}{m}\sum_{i=1}^{m}\|z_i-\tilde{z}_i\|^2 \tag{9}$$

其中：$w(t)$ 表示的是在训练过程中正则项对鉴别器贡献是动态的，训练初期，由于网络参数较差，该正则项的贡献为 0，随着训练的进行，贡献程度逐步上升，最后趋于一个稳定值。接下来讨论 SENet 模块如何提取任务相关的特征以及如何嵌入到网络中训练的。图 2 是 SENet 模块的结构图。

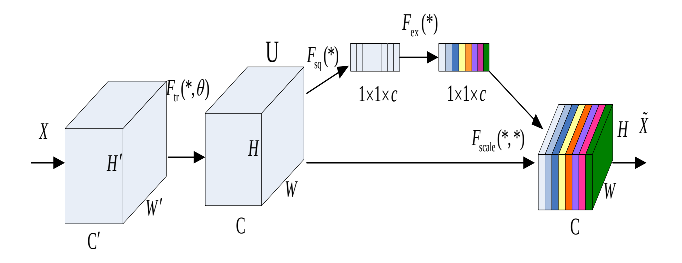


图 2　SENet 模块结构

Fig. 2　Structure of senet Model

给定一个输入 X，$X\in\mathbb{R}^{W'\times H'\times C'}$，通过一系列卷积映射变换 $F_{tr}(*,\theta)$，得到 C 个特征图的集合 U，$U=[u_1,u_2,\cdots,u_C]$，其中 $u_i\in\mathbb{R}^{W\times H}$。下面通过三个操作来重标定经过映射得到的特征 U。

a）压缩 $F_{sq}(*)$ 操作：为了探索特征通道相互之间的依赖性，首先考虑学习特征图中每一个通道的信息。由于卷积核是在一个局部感受野进行操作，因此，卷积变换之后得到的节点都无法获取该局部感受野以外的信息。压缩操作将全局空间信息压缩到一个通道描述符中，通过全局平均池化来生成描述特征通道信息的统计量 z，统计量的第 c 个元素计算方式如下：

$$z_c=F_{sq}(u_c)=\frac{1}{W\times H}\sum_{i=1}^{W}\sum_{j=1}^{H}u_c(i,j) \tag{10}$$

其中; $z\in\mathbb{R}^{C}$，代表着整个图像的信息。

b）激活 $F_{ex}(*)$ 操作：它是一个类似于循环神经网络中门的机制。由两个 全连接层组成，将特征维度降低到输入的 1/16，然后经过 ReLu 激活后，通过一个全连接层升到原来的维度，最后通过一个 Sigmoid 的门获得 0~1 之间归一化的权重。具体表达式如下：

$$s=F_{ex}(z,W)=\sigma(g(z,W))=\sigma(W_2\delta(W_1 z)) \tag{11}$$

其中: δ 是指 ReLu 激活函数，σ 是指 Sigmoid 函数，$W_1\in\mathbb{R}^{\frac{C}{r}\times C}$，$W_2\in\mathbb{R}^{C\times\frac{C}{r}}$。参数 W 用来显式地建模特征通道间的相关性。

c）重标定 $F_{scale}(*)$ 操作：将激活操作得到的权重当作是特征选择之后的特征通道的重要性指标，然后与之前的通道逐个加权相乘，完成在通道维度上的对原始特征的重标定。对第 c 个特征通道的加权公式如下：

$$\tilde{x}_c=F_{scale}(u_c,s_c)=s_c\cdot u_c \tag{12}$$

其中：$\tilde{X}=[\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_c]$，$S=[s_1,s_2,\cdots,s_c]$。

2.2　算法流程

基于特征重标定的生成对抗网络半监督分类算法流程如下所示，$f_\theta(x)$ 表示输入数据随机预处理函数；L 表示带标签的数据集合；B 表示每一个批大小 Batchsize 组成的数据集合；$w(t)$ 表示无监督权重的动态时间函数；用 Xavier 方法进行参数初始化。

初始化：批大小 Batchsize：m=100，即每一次参数更新时所需的样本数；iterations：批迭代次数，遍历一次训练集所需 m 的个数，即总的样本数除以 m 的值；epoch：一个 epoch 等于遍历训练集一次。设超参数 k=1，即训练鉴别器 k 次才训练生成器 1 次；噪声样本为服从 $z\sim U[-1,1]$ 的 100 维向量。

for number of training iterations do

for k steps do

抽样 m 个噪声样本{ $z^{(1)}$，$z^{(2)}$ …… $z^{(m)}$ }

抽样 m 个数据样本{ $x^{(1)}$，$x^{(2)}$ …… $x^{(m)}$ }

计算鉴别器的对抗损失：

$$\text{loss}_{adv}=-\frac{1}{m}\sum_{i=1}^{m}[\log D(x^{(i)})+\log(1-D(G(z^{(i)})))]$$

计算不同随机状态下鉴别器对真实样本的输出：

$z_i\in B\leftarrow D_\theta(f(x_i\in B))$

$\tilde{z}_i\in B\leftarrow \tilde{D}_\theta(f(x_i\in B))$

计算鉴别器的监督损失和无监督损失

$loss_{label}=-\frac{1}{|B|}\sum_{i\in(B\cap L)}\log(z_i[y_i])+w(t)\frac{1}{|B|}\sum_{i\in B}\|z_i-\tilde{z}_i\|^2$ 根据损失计算鉴别器的梯度：

$grad(\theta_d)=\nabla\theta_d[loss_{adv}+loss_{label}]$

根据梯度更新鉴别器的参数：

θ_d=$Adam(\text{grad}\theta_d,\theta_d)$

end for

抽样 m 个噪声样本{ $z^{(1)}$，$z^{(2)}$ …… $z^{(m)}$ }

计算生成器的对抗损失

$\text{loss}_{gen}=E_{z\sim p_{noise}}\log[-D(G(z))]$

根据损失计算梯度并更新生成器的参数：

$grad(\theta_g)=\nabla\theta_g[loss_{gen}]$

θ_g=$Adam(\text{grad}\theta_g,\theta_g)$

end for

3　实验结果及分析

本章实验选取了 SVHN[14]数据集进行实验验证。实验是在 Theano[15]深度学习框架上实现，在单块 GPU 型号为 GTX1080 上运行。本实验的生成器以 DCGAN 生成器为基准，鉴别器引入 SENet 模块来增加模型的非线性。初始学习率设为 0.0003，模型中同样使用 WN、BN 以及 dropout 策略。关于网络模型的参数配置如表 1 所示。

表 1　SVHN 数据集的网络参数配置

Table 1　Network parameter configuration of the SVHN data set

鉴别器 D	生成器 G
输入 32×32 三通道 类标签 $y\in R^{10}$	输入 Noise $\in R^{100}$
	MLP 8192 units
Dropout=0.2	ReLU，BN
3×3 conv，64，lReLU，WN	Reshape 512×4×4
3×3 conv，64，lReLU，WN	5×5 deconv，256，stride 2
3×3 conv，64，lReLU，WN	ReLU，BN
Dropout=0.2	
3×3 conv，128，lReLU，WN	5×5 deconv，128，stride 2
3×3 conv，128，lReLU，WN	ReLU，BN
3×3 conv，128，lReLU，WN	
SENet block	
Dropout=0.2	
3×3 conv，256，lReLU，WN	
NIN，128，lReLU，WN	5×5 deconv，3，stride 2,
NIN，128，lReLU，WN	Tanh，WN
Global pool layer	
Denselayer 10 Units with WN	

为了对比加入模块前后的生成样本质量和半监督分类精度，保证其他模型的配置相同。在生成样本的质量方面，FR-GAN 与无 SENet 模块仅有正则项的 GAN 作比较，视觉对比结果如图 3 所示。使用 IS[7](Inception Score)指标进行定量对比，对比结果如表 2 所示。



真实样本

Real sample



正则项 GAN

GAN with regular term



本文 GAN

FR-GAN

图 3　SVHN 数据集上生成样本对比图

Fig. 3　Comparison of Generative Images in SVHN Datasets

表 2　不同模型生成样本的初始得分

Table 2　Initial scores of samples generated by different models

方法	真实样本	GAN without SENet	GAN without regular term	FR-GAN
score±std.	3.24±0.17	2.89±0.02	3.02±0.14	2.95±0.43

在半监督学习方面，分别设每类带标签的样本数 50、100，带标签的训练样本为 500、1000。如表 3 所示，与传统的半监督学习算法相比，除文献[5]中算法，均比其他算法较好；文献 5 仅针对单一的分类任务，分类性能较好，FR-GAN 虽在图像分类任务较次，但其损失策略能够帮助其在图像生成任务比常规 GAN 较好。

当使用所有标签样本进行训练时，将部分正确标签样本给予错误标签时，所提模型能够减小分类精度的下降程度，使其不会出现太大变化，即提高网络的容错能力。

如图 4 所示，（a）表示的是常规 GAN 在一定比例的错误标签下的分类精度变化，随着比例的增加，分类精度下降较快；（b）表示 FR-GAN 在一定比例的错误标签下的分类精度变化，可以看出，在错误标签比例占 50%时，精度基本保持在一定范围。

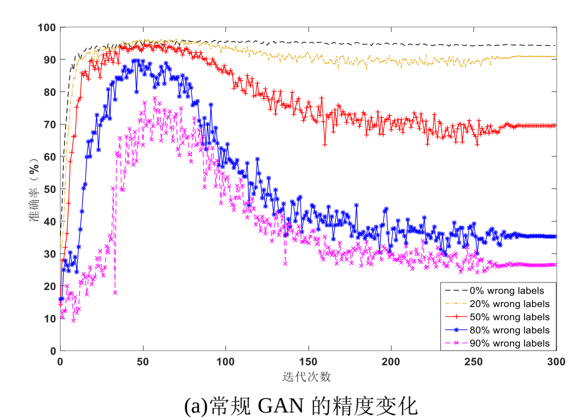
表 3　SVHN 数据集测试错误率对比结果

Table 3　SVHN data set test error rate comparison results

方法	测试错误率（%）#标签数	
	500	1000
M1+M2[1]	-	36.02 ± 0.10
VAT[6]	-	24.63
ADGM[2]	-	22.86
SDGM[2]	-	16.61 ± 0.24
Improved-GAN[7]	18.44 ± 4.8	8.11 ± 1.3
ALI[10]	-	7.42 ± 0.65
-model[5]	7.05 ± 0.30	5.43 ± 0.25
Temporal ensembling[5]	5.12 ± 0.13	4.42 ± 0.16
FR-GAN with regular term	15.62 ± 0.16	8.13 ± 0.22
FR-GAN	13.41 ± 0.20	7.20 ± 0.16

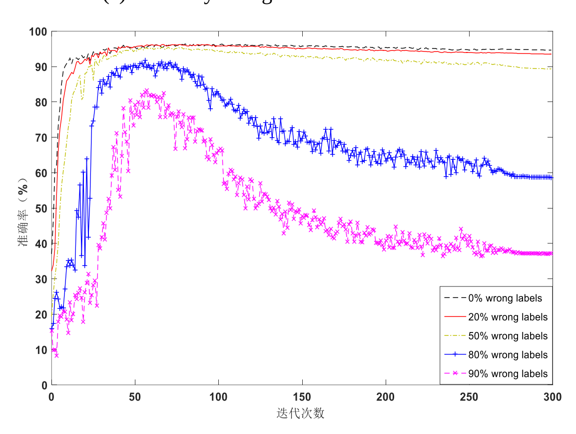
4　结束语

本文主要针对传统鉴别器的训练模式难以提取到更抽象的和任务相关的鲁棒性特征，提出了基于特征重标定的生成对抗网络，为了学习到任务相关的特征，为鉴别器引入模型在不同状态下的无监督特征级均方差损失，对真实样本中的同一输入得到的不同输出进行惩罚。与此同时，传统的鉴别器的结构是卷积加池化模块组成的，本文在鉴别器的中间层引入 SENet 模块，自动学习每一个特征通道的重要程度，提取任务相关有用的特征抑制任务无关的特征，从而提高半监督学习表现。



(a)常规 GAN 的精度变化

(a) Accuracy change of conventional GAN



(b)FR-GAN 的精度变化

(b) FR-GAN accuracy variation

图 4　错误标签下的分类精度变化对比

Fig. 4　Comparision of Classification Accuracy Rate in Wrong Labels Case

参考文献：

[1] Kingma D P, Rezende D J, Mohamed S, et al. Semi-supervised learning with deep generative models [EB/OL]. (2014-10-31).https://arxiv.org/pdf/1406.5298.pdf.

[2] Maaløe L, Sønderby C K, Sønderby S K, et al. Auxiliary deep generative models [EB/OL]. (2016-06-17). https://arxiv.org/pdf/1602.05473.pdf.

[3] Li Chongxuan, Zhu Jun, Zhang Bo. Max-margin deep generative models for (semi-) supervised learning [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2016, 3(9): 1-10.

[4] Rasmus A, Valpola H, Honkala M, et al. Semi-supervised learning with ladder networks [J]. Computer Science, 2015, 9: 1-9.

[5] Laine S, Aila T. Temporal ensembling for semi-supervised learning [EB/OL]. (2017-03-15). https://arxiv.org/abs/1610.02242.

[6] Miyato T, Maeda S, Koyama M, et al. Distributional smoothing with virtual adversarial training [J]. Computer Science, 2016, 5(2): 72-86.

[7] Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training gans [EB/OL].(2016-06-14). https://arxiv.org/pdf/1606.03498.pdf.

[8] Springenberg J T. Unsupervised and semi-supervised learning with categorical generative adversarial networks [J]. Computer Science,

2015, 6 (5): 34-42.

[9] Li Chongxuan, Xu Kun, Zhu Jun, et al. Triple generative adversarial nets [EB?OL]. https://arxiv.org/abs/1703.02291.

[10] Dumoulin V, Belghazi I, Poole B, et al. Adversarially learned inference [EB?OL]. https://arxiv.org/abs/1606.00704..

[11] Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks [C]// Advances in Neural Information Processing Systems. 2015: 2017-2025.

[12] Jie Hu, Li Shen, Gang Sun. Squeeze-and-Excitation Networks [C]//Proc of. Computer Vision and Pattern Recognition. 2017: 1-8.

[13] Zeiler M D, Krishnan D, Taylor G W, et al. Deconvolutional networks [C]// Computer Vision and Pattern Recognition, 2010: 2528-2535.

[14] Netzer Y, Wang Tao, Coates A, et al. Reading digits in natural images with unsupervised feature learning [C]//Proc of NIPS Workshop on Deep Learning and Unsupervised Feature Learning. 2011: 5-13.

[15] Bastien F, Lamblin P, Pascanu R, et al. Theano: new features and speed improvements [J]. Computer Science, 2012, 8 (1): 51-63.